SANAS

Un disque de 30'000 Go à l'EPFL et une utilisation dans la Faculté STI

Aristide.Boisseau@epfl.ch, Domaine IT & Laurent.Klinq@epfl.ch, Faculté STI

Cet article est dédié à la mémoire de Paul Debefve, coordinateur informatique de la faculté STI, décédé subitement le mardi 10 mai 2005

Historique

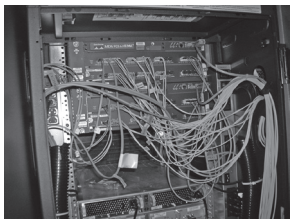
Un projet de consolidation du stockage a démarré au milieu de l'année 2004 sous le mandat de la DIS. Le but de ce projet appelé SANAS est de proposer de la capacité de stockage réseau brut SAN (Storage Area Network) ainsi qu'un service de *fichiers* de type NAS *(Network Attached Storage)* en complément du service AFS existant. La solution retenue est celle de la compagnie EMC[2] (voir article **Serveurs centraux, de calculs et de fichiers: nouveautés pour 2005** paru dans le FI10/04). En pratique, cela représente 30 To visibles pour les utilisateurs (environ 20 To pour le service NAS et 10 To pour les clients SAN), avec une capacité identique pour le site de sauvegarde, soit 60 To au total.

Architecture d'un SAN-NAS

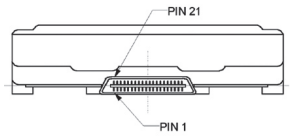
Voir la configuration de SAN-NAS sur l'illustration en page suivante.

Un réseau de stockage, Storage Area Network, SAN

Malgré l'adage connu qu'un dessin représente mille mots, il semble nécessaire d'expliquer par le détail cette architecture:

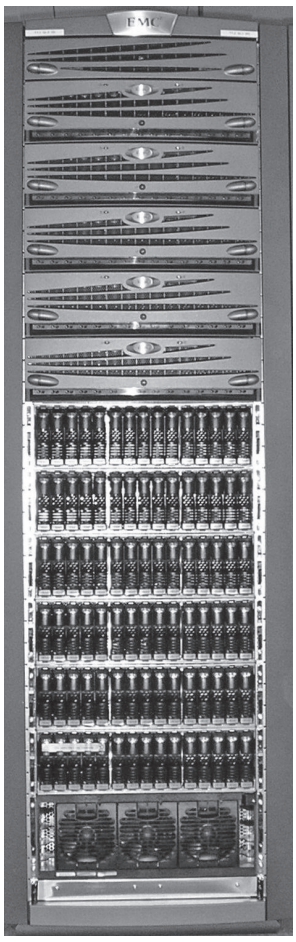


Un média unique pour le transfert des données, la technologie fibre autorise un débit de 2 Go par seconde.

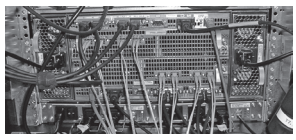


Au départ, un disque dur haute performance disposant de deux interfaces fibres, actuellement d'une capacité de 146 Go, la nouvelle génération de disque de 300 Go va bientôt être certifiée sur cet équipement.

Celui-ci prend place en compagnie de 14 collègues pour remplir une unité de rack d'une hauteur de 3u (13.34 cm) dénommée *Enclosure* (en anglais dans le texte).



Chaque *Enclosure* est connectée par deux boucles fibre à un contrôleur. Dans l'équipement utilisé dans la configuration de l'École, il se dénomme CLARiiON C700.
Un CLARiiON C700 est équipé de 4 processeurs à 3 GHz, 8 Go de mémoire, il peut gérer jusqu'à 240 disques.
Ceci représente une baie de stockage. Le C700 est localisé en bas du rack, vu de face on voit les ventilateurs, à l'arrière on voit la connectique.

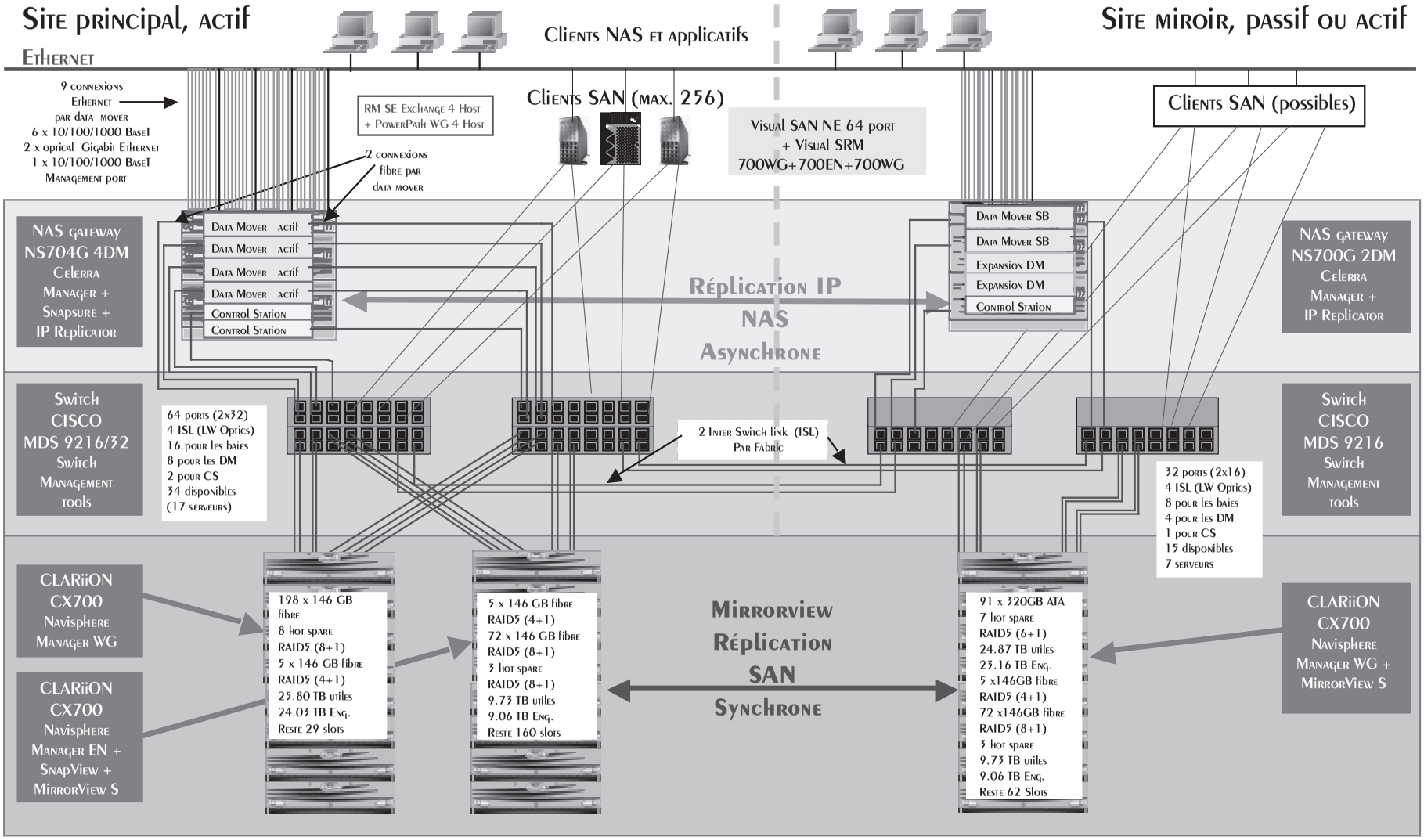


Pour assurer une fiabilité optimale, chaque groupe de 9 disques est configuré en RAID 5 (8+1), un RaidGroup dans la terminologie idoine. Pour le NAS, chaque RaidGroup est combiné avec trois systèmes équivalents qui sont divisés en segments de 250 Go physiques. Ainsi, les données sont réparties sur 36 disques assurant des performances élevées, on parle de 36 axes dans le jargon (technique des stripes et des slices). Chaque baie de stockage (CLARiiON C700) est connectée à deux commutateurs fibres (Cisco 9232, pour la redondance et le *load balancing*) qui desservent les clients. Pour séparer les communications entre chaque client, une technique similaire au masque d'adressage IP est utilisée, qui s'appelle Zoning et LunMasking.

Cette architecture est complexe et on pourrait se poser la question de sa pertinence. En pratique le matériel n'est que la fondation d'un SAN. Le logiciel de gestion représente la clé de voûte de l'ensemble.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

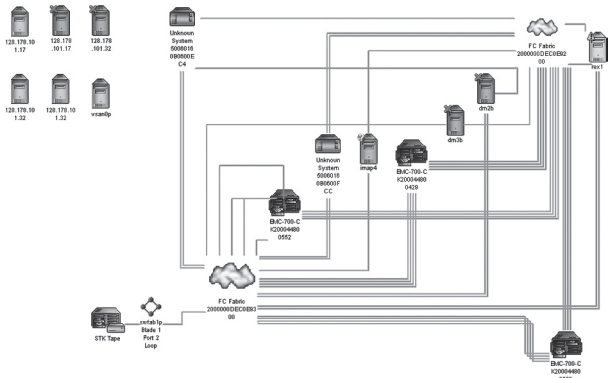
Schéma fonctionnel



Version 08 du 26 Novembre 2004

fig. 1 – architecture de SAN-NAS; avec du haut en bas, les clients sur le réseau, les serveurs clients du SAN, les serveurs du NAS, les commutateurs fibres et finalement les baies de stockage

Par une interface Web réalisée par des applets Java, on accède à la gestion de l'ensemble des fonctionnalités par une interface unique.



Pour les aficionados de la ligne de commande, il est également possible d'accéder à des fonctions plus spécifiques réservées aux initiés. Dans notre configuration, deux baies de stockage similaires ont été mises en place sur le site de production, représentant logiquement les fonctionnalités SAN et NAS.

Les clients (Serveurs Windows, Solaris, Linux, interfaces NAS) sont connectés par des liens fibres. Pour assurer une redondance élevée, chaque client possède deux liens fibres au minimum. Dans le jargon, un lien fibre est dénommé *une patte*, rendant l'apparition d'un mille-pattes peu plausible.

Une sécurité très élevée, deux sites

Pour assurer le minimum de perturbations en production, un site *miroir* existe en parallèle du site de production du DIT. Le miroir est situé dans le nouveau bâtiment de la Faculté IC distant de ~1'200 m. Actuellement en mode passif, il est configuré pour récupérer des données ou assurer en mode dégradé la reprise provisoire du site de production.

Un mode site de production actif - site miroir actif peut être envisagé dans une utilisation future.

La configuration plus simple du site miroir se compose d'une baie de stockage (CLARiiON C700) , de deux commutateurs fibre (Cisco MDS 9216) et de deux *data mover* (Celerra NS702G).

Pour assurer un service sans dégradation pour les clients SAN, une copie synchrone est assurée avec des disques fibres pour une capacité de 10 To. La copie asynchrone des données pour le NAS s'effectue par des connexions Ethernet dédiées. Pour réduire le coût, des disques SATA de 320 Go sont utilisés.

Une copie de sauvegarde sur bande

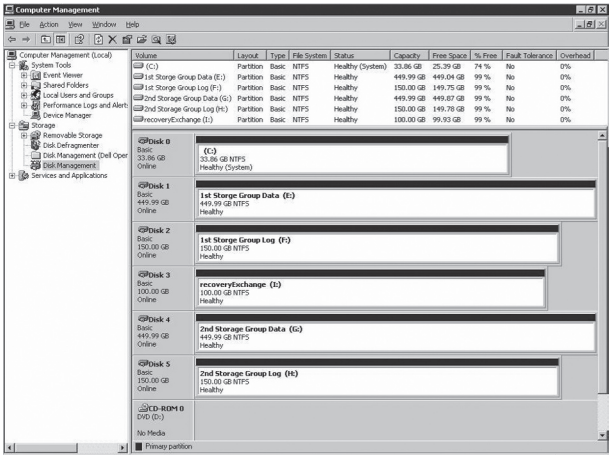
La protection de l'ensemble des données est réalisée par le robot STK du DIT déjà utilisé pour le service de sauvegarde centralisée. Pour le NAS, un lecteur de bandes est directement intégré aux commutateurs fibres et offre avec le protocole NDMP des performances élevées, de 15 à 25 Mo par seconde. Il est rattaché à une des têtes NAS du site miroir. Les données du NAS étant répliquées, les sauvegardes pourront s'effectuer depuis le site miroir. Cette architecture permet de travailler 24h/24h 365j/365j sans perte de performance sur le site de production.

Pour le SAN, un client Veritas NetBackup devra être installé et configuré pour chaque client.

Services SAN

Comme indiqué précédemment, les clients sont connectés par au moins une patte, une connexion fibre entre le commutateur et lui-même, la carte qui remplit ce rôle est dénommée un HBA. Cette connexion physique entraîne une proximité relative (moins de 100m) pour l'équipement.

Une fois la carte (HBA) et les pilotes installés, l'ordinateur voit l'espace mis à disposition comme un disque dur résidant dans son équipement, alors que celui-ci réside dans la baie de stockage.



Les performances très élevées de la baie permettent la création rapide de clone d'un disque dur SAN avec un délai sans équivalence. Si la synchronisation initiale prend des heures (tâches de fond), la mise à jour de clone prend par la suite quelques minutes.

L'étape ultime serait de supprimer les disques durs locaux des ordinateurs clients et de démarrer le système à partir d'un disque résidant sur le SAN.

L'installation d'agent spécifique à un logiciel métier, pour une base de données ou un serveur de messagerie permet d'étendre cette fonctionnalité de sauvegarde à un élément logique.

Dans notre cas, elle sera utilisée pour sauvegarder régulièrement l'état des boîtes aux lettres électroniques d'Exchange.

Dans la configuration de l'École, l'espace SAN représente 30% de l'espace total, soit 10 To. Les premiers clients du SAN seront les serveurs de messagerie de l'École, IMAP sous Unix Solaris et Exchange sous Windows 2003.

Une consolidation de serveurs, un Network Attached Storage, NAS

En pratique, un service NAS est lui-même client du SAN. Dans notre cas, une baie de stockage est dédiée au NAS, mais on peut envisager une configuration mixte. Dans la configuration de l'École, l'espace NAS représente 70 % de l'espace total, soit 20 To.

Connecté par 4 pattes à chacun des commutateurs, le NAS est constitué de 4 serveurs de déplacement de données *DataMover* (Celerra NS704G) et d'une console de contrôle en production.

Chaque serveur de déplacement de données, Celerra NS704G, possède 4 Go de mémoire et 2 processeurs à 3 GHz.

Contrairement à un serveur généraliste, un DataMover ne sait que transférer des données depuis le SAN vers des clients connectés par Ethernet. Il possède 8 ports Ethernet de 1 Gb/s.

Un serveur *virtuel*

Ce serveur virtuel possède tous les attributs d'une machine réelle. En particulier un nom Wins et DNS, une adresse IP, et même une adresse MAC programmable.

Représenté par serveur virtuel sur le NAS, *Virtuel Data Mover, VDM,* il réside sur un *DataMover* et peut automatiquement se transférer sur un autre *DataMover* en cas de problème ou d'intervention de maintenance.

Ce VDM contient des espaces de données *(file system)* qui peuvent atteindre 2 To actuellement, 16 To dans une prochaine version.

L'avantage immédiat d'un NAS est la possibilité d'augmenter la taille d'un file system à tous moments, cela sans interrompre son utilisation et avec une rapidité déconcertante (moins de 5 minutes pour passer de 500 Go à 1000 Go).

Il existe la possibilité de travailler indifféremment avec des partages NFS (Unix) ou CIFS (Windows, Samba).

Actuellement, le système supporte NFS version 3, ce qui implique une limitation d'accès par adresse IP.

Le support prévu de NFS version 4 permettra d'intégrer les fonctionnalités d'identification à la connexion et l'utilisation de liste de droits (ACL).

Une sauvegarde de l'état des fichiers à un instant donné

La capacité de sauvegarder à intervalles réguliers l'état des données, *snapshot* sur un espace dédié, est également une fonctionnalité particulièrement intéressante.

Son utilisation se résume simplement à taper **.ckpt** dans n'importe quel répertoire, sous n'importe quels clients, pour obtenir la vision des fichiers à un instant donné.

Ceci est également possible avec le client *Windows shadow copy client* (VSS).

Cette fonctionnalité accessible par chacun représente une avancée majeure, évitant souvent le recours à une récupération depuis une bande.

Une limitation de l'utilisation de l'espace par Quota

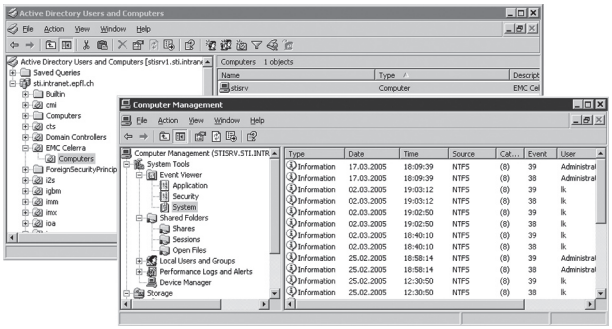
Pour permettre une utilisation rationnelle de l'espace disque mis à disposition d'un groupe d'usagers, la présence d'un gestionnaire de quota permet un contrôle de l'espace disque utilisé.

Des données exemptes de virus avec l'accès CIFS

Pour éviter la contamination des données, un antivirus McAfee empêche l'entrée des fichiers infectés. Malheureusement, cette fonction n'est pas présente avec le protocole d'accès NFS.

Une intégration élevée avec Windows

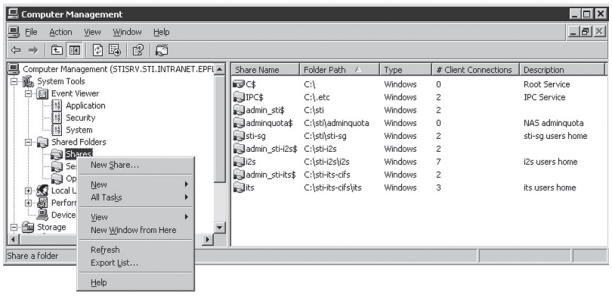
La beauté de l'implémentation de EMC² est de donner aux administrateurs Windows la visualisation d'un serveur sous une forme similaire à une machine réelle:



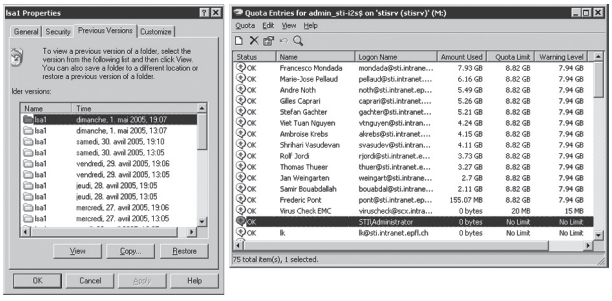
une console d'administration MMC

On retrouve ainsi l'ensemble des outils de gestion d'un serveur PC classique.

Sur un *filesystem*, on crée des partages de fichiers avec la console d'administration.



L'intégration des fonctionnalités des Snapshots et des Quotas est très élevée, l'utilisation des API de Microsoft, permettant un affichage identique:



L'apparence est trompeuse, puisque la tête NAS fonctionne sous un système d'exploitation à base d'Unix BSD avec un format de données spécifiques qui ajoute le SID de Windows aux UID et GID habituels d'UNIX.

Il est important d'en avoir conscience dès que l'on aborde son intégration dans un environnement hétérogène.

La mise en œuvre pour la Faculté STI

Une vision sur le long terme

Fin 2003, un espace total de 6 To est disponible dans les serveurs de la Faculté. Cet espace est hébergé sur des serveurs de capacité et de technologie fort variés.

Dans l'optique du remplacement de serveurs obsolètes ou dont la maintenance sera échue, une demande hors enveloppe d'un NAS a été soumise en 2004. Cette demande a été incluse dans le projet SANAS conjointement à une demande similaire issue de la Faculté SB.

Paul Debefve et Laurent Kling ont suivi l'évaluation et le choix de l'équipement. La Faculté STI étant choisie pour réaliser le pilote de l'implémentation avec un partage commun CIFS et NFS.

Un espace commun pour les accès NFS et CIFS

L'idée d'un espace commun accessible par les deux protocoles de transfert les plus utilisés dans la faculté STI vient immédiatement à l'esprit, particulièrement quand la plateforme technique utilisée se base sur une mouture d'Unix.

Il est nécessaire de dépasser les interfaces graphiques, voir même le mode ligne de commande pour pouvoir mettre en place un tel système et il faut disposer de:

- un mécanisme automatisé de synchronisation, Adldap^2;
- la capacité de celui-ci de s'adapter rapidement aux besoins;
- la présence d'un expert dont la création de script personnalisé résout les problèmes non imaginés par les concepteurs du NAS: Daniel Meier d'EMC^2;
- l'administrateur du NAS à l'EPFL: Aristide Boisseau du DIT;
- un administrateur éclectique aimant marier la carpe et le lapin: Laurent Kling de la STI;
- un certain temps, comme aurait dit Fernand Renaud, pour tester les différentes stratégies.

Un DataMover dans le détail

Le serveur de transfert de données est un ordinateur qui exécute un système d'exploitation Unix BSD avec un système de fichiers propriétaire.

Pour comprendre les mécanismes d'identification entre l'ancien et le Nouveau Monde, un bref rappel:

	Unix	**Windows**
Usager	Le nom de l'usager est mis en relation avec son User Identification (UID). Un objet sur un disque possède toujours un UID. L'UID est unique	L'unique identifiant est le System Identification (SID), il est attribué à chaque objet. Dans un domaine Active Directory, tous les SID ont une base commune. Le début d'un SID est semblable à tous les ordinateurs Windows, il est issu d'une désignation X500 (OID). Un SID ne peut pas être normalement recréé. Il est unique dans la durée.
Groupe	L'usager est membre de groupes de sécurité, il possède un groupe préférentiel. Chaque groupe de sécurité possède un Group Identification (GID). Le GID est unique	

Un DataMover possède deux mécanismes d'identification avec Active Directory.

UserMapper	**SecurityMapper**
Le mécanisme standard d'authentification, dès l'introduction d'un serveur *virtuel* dans un domaine, récupère l'ensemble des groupes auquel appartient l'usager utilisé, puis est capable d'établir une relation entre un SID de l'usager et un UID à la volée, de même entre le SID des groupes de sécurité et leur GID. Les UID et GID sont générés automatiquement.	Un mécanisme interne de gestion similaire à Unix, utilisant deux tables pour les UID et GID. Un mécanisme de découverte et de résolution entre le SID et le nom des usagers ou du groupe de sécurité étant réalisé à la volée par un mécanisme interne non documenté.
Import-export: il est uniquement capable d'importer des données d'usager et de groupe de sécurité avec leur SID	Import-export: on peut forcer les tables UID et GID.

Le projet avec **Adldap^2**:

	Adldap^2	NAS SecurityMapper
1	Générer les tables UID et GID pour introduire le *serveur virtuel NAS* dans le domaine. En conséquence, les opérations qui suivent doivent être réalisées.	Malgré un succès d'estime rapide, il est vite apparu que l'intégration est plus complexe. Après une observation du fonctionnement normal (avec UserMapper), il apparaît que la quasi-totalité des groupes et des comptes du domaine Active Directory doivent êtres présents dans les tables du SecurityMapper
2	Vérifier la conformité de la nomenclature des usagers et des groupes de sécurité du domaine AD selon le RFC 1123.	Pour éviter la présence de caractères indésirables.
3	Créer la table des usagers UID. Établir la relation biunivoque entre l'usager AD et son UID provenant du LDAP.	Remplacer la table de correspondance usager = UID. La correspondance SID = UID est réalisée par un mécanisme interne.
4	Créer la table des groupes GID. A: établir la relation biunivoque entre le groupe de sécurité unité-StaffG AD et son GID provenant du LDAP. B: créer et conserver une table de relation biunivoque entre tous les autres groupes de sécurité AD et un GID générés à la volée.	Remplacer la table de correspondance groupe de sécurité = GID. La correspondance SID = GID est réalisée par un mécanisme interne.

5	Mettre en place un mécanisme pour mettre à jour régulièrement le SecurityMapper dans le DataMover

Cette démarche s'est réalisée sans problèmes particuliers. Un écueil majeur est cependant apparu dans le cas de modification du nom d'un objet, car le mécanisme de résolution interne du SecurityMapper entre le SID et l'UID ou le GID ne possède pas d'interface de programmation pour interagir avec lui.

L'utilisation alternative du UserMapper serait possible, mais le cas suivant peut générer des fichiers incohérents sur le DataMover.

- création d'un compte dans Active Directory avec la génération automatique d'un SID pour l'usager;
- utilisation de celui-ci sur le NAS;
  - résolution du SID de l'usager et création d'un UID[1] à la volée par le NAS;
  - création de fichier par l'usager avec l'UID[1];
- remplacement périodique des relations entre SID et UID[2] défini par LDAP par **Adldap**[2];
  - perte de l'accès au fichier, car l'UID[1] généré à la volée ne correspond pas à l'UID[2] contenu dans le LDAP.

En conclusion, dans la version actuelle du NAS, il n'est pas possible de garantir un fonctionnement sans la possibilité de cas limites avec l'introduction de données externes. En particulier, l'absence d'interface de programmation sur les mécanismes internes du NAS qui permettrait de résoudre la relation entre SID et UID-GID. Il est possible que ces limites disparaissent avec une prochaine version du code du NAS.

Un espace NFS séparé

L'objectif d'un espace commun ne pouvant se concrétiser, le protocole CIFS semble être celui qui offre le moins d'inconvénients.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Faculté STI Sciences et Techniques de l'ingénieur

adldap2: Tableau de bord

Rapport du 06.05.2005 14.02 — 06 05 2005 11:57 — Administrateur Laurent.kling@epfl.ch

Active Directory		EPFL
Unité		
	Existante	125
	A créer	0
	OU Services	284
Usager		
	Existant	1'157
	A créer	0
	A déplacer	0
	A archiver	0
	Archiver, mais utiliser	0
	Archiver	429
	Autres domaines	40
	Service	249
	Autres	154

NAS				
Nb		Attribué	Utilisé	
15	12.0%	933.00 Go	72.39 Go	7.8%
		Usager	Commun	
166	14.3%	72.39 Go	0.00 Go	

Tableau de bord d'Adldap[2] récapitulatif

ManagedByMail	UnitGID	Name	NameStri	Path	UseNAS	NA	NAS path	NAS userQuota	NAS adminQuota	NAS allocTotal	NAS allocCommon	NAS usedUser	NAS usedC
laurent.kling@ epfl.ch	20060	dc=sti	dc=sti	dc=sti,dc=	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20061	sti-dec	stidec	ou=sti-dec	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10380	sti-do	stido	ou=sti-do,(	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20063	sti-sg	stisg	ou=sti-sg,(	VRAI	L	L:	2899102924	16106127360	16106127360	16106127360	0	0
laurent.kling@ epfl.ch	10381	sti-ge	stige	ou=sti-ge,(	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10382	sti-rh	stirh	ou=sti-rh,c	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10383	sti-fi	stifi	ou=sti-fi,ou	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10384	sti-re	stire	ou=sti-re,c	FAUX			0	0	0	0	0	0
jean-pierre.moinat@ epfl.ch	10385	sti-it	stiit	ou=sti-it,ou	VRAI	L	L:sti-sg	4831838208	17716740096	61203283968	17716740096	1558036480	0
jean-pierre.moinat@ epfl.ch	10386	sti-in	stiin	ou=sti-in,o	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10388	sti-sc	stisc	ou=sti-sc,(	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20065	sti-sel	stisel	ou=sti-sel,	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10406	sel-ge	selge	ou=sel-ge,	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	14804	sel-ens	selens	ou=sel-ens	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20066	sti-sgm	stisgm	ou=sti-sgm	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10306	sgm-ge	sgmge	ou=sgm-g(	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	14805	sgm-ens	sgmens	ou=sgm-en	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20067	sti-smx	stismx	ou=sti-smx	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10806	smx-ge	smxge	ou=smx-g(	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	14806	smx-ens	smxens	ou=smx-en	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20070	sti-smt	stismt	ou=sti-smt	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	11006	smt-ge	smtge	ou=smt-ge	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	14807	smt-ens	smtens	ou=smt-en	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20072	sti-sgb	stisgb	ou=sti-sgb	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	11007	sgb-ge	sgbge	ou=sgb-ge	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	14808	sgb-ens	sgbens	ou=sgb-en	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	20042	ise	ise	ou=ise,dc=	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10311	ise-ge	isege	ou=ise-ge,	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10420	lei	lei	ou=lei,ou=	FAUX			0	0	0	0	0	0
laurent.kling@ epfl.ch	10312	leni	leni	ou=leni,ou	FAUX			0	0	0	0	0	0

Tableau de bord d'Adldap[2]: un exemple d'état parmi quatorze disponibles, l'utilisation du NAS par unité

Les unités qui le désirent peuvent obtenir un espace NFS au détriment d'un espace CIFS.

Une synchronisation LDAP - Active Directory - NAS = Adldap²

Dans la pratique, il faut gérer la création d'unité, la création et le déplacement d'usager, la gestion des droits de sécurité, et automatiser ce travail.

Ce constat a généré le développement d'**Adldap**² qui intègre toutes les fonctionnalités de Adldap avec l'ajout de la gestion du NAS. **Adldap**² est utilisé actuellement dans trois Facultés par l'intermédiaire de leurs responsables de Domaine AD:

- Sourythep Samoutphonh pour I&C,
- Yvon Roucher pour SV
- et Laurent Kling pour STI.

Un article lui a été consacré dans le FI 3/04, ***http://dit.epfl.ch/publications-spip/article.php3?id_article=512***.

Un espace de 3 Go pour chaque collaborateur

La faculté STI comprend 1'157 personnes (état du 6 mai 2005) réparties dans 125 structures avec 125 comptes d'administration locaux (voir figure en page précédente).

Ces données sont issues du tableau de bord inclus dans **Adldap**² qui permet de visualiser l'état complet du domaine. Les données spécifiques à un état sont représentées par des vues supplémentaires. La possibilité de suivre les modifications sur la durée permet de comprendre l'évolution de la population concernée.

125 gestions personnalisées

À l'écoute des demandes des usagers, une gestion centralisée peut vite apparaître insurmontable à mettre en place pour faire face aux demandes contradictoires sur l'utilisation d'un serveur de fichier. Certains laboratoires désirent que l'espace soit alloué personnellement avec un accès strictement privé, d'autres laboratoires veulent bannir les espaces privés et disposer uniquement d'un lieu commun accessible à tous. La nécessité d'un canevas commun personnalisé pour chacune des unités concernées (laboratoires, instituts, administration, ateliers), apparaît rapidement. Les limites de la configuration du NAS sont définies par des contraintes techniques:

- le débit d'une sauvegarde qui n'excède pas 100 Go/heure
- la durée de reconstruction d'un RAID.

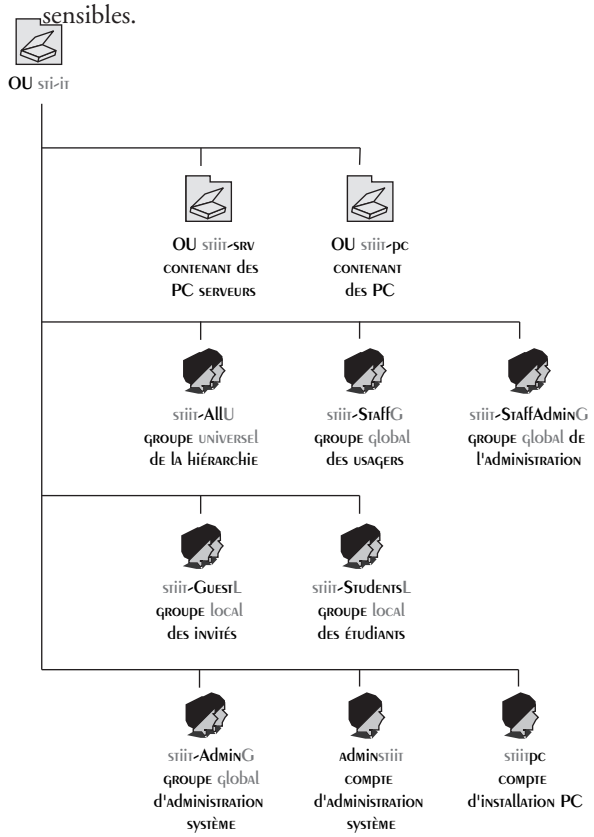
La réponse se décompose en quatre éléments:

1. une structure unifiée
2. une configuration personnalisée
3. des quotas modulables
4. une gestion indépendante d'une plate-forme.

Réaliser une structure unifiée

Forte de l'expérience précédente réalisée dans le domaine Active Directory STI, la structure proposée correspond aux demandes des usagers. Deux nouveaux groupes de sécurité ont été introduits:

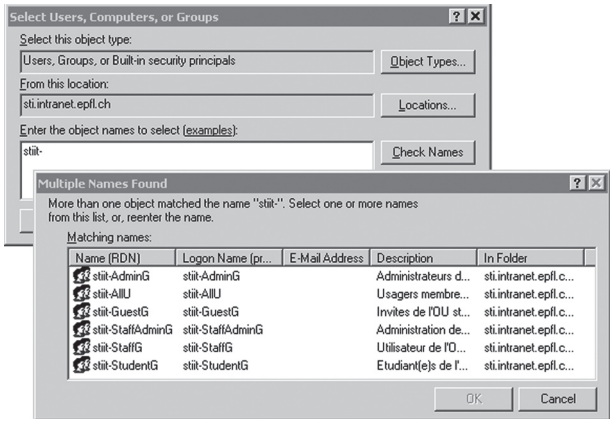
- un groupe universel unité-AllU qui contient l'ensemble des collaborateurs sous le niveau d'un laboratoire, institut ou Faculté.
- un groupe de sécurité unité-StaffAdminG qui possède son corollaire dans le NAS pour héberger les données sensibles.



La composition du nom des groupes de sécurité permet une efficacité optimale, il se décompose en quatre parties:

1. le préfixe correspondant au nom de l'unité sans «-» pour éviter des confusions
2. un séparateur «-»
3. la fonction du groupe de sécurité
4. un suffixe qui représente la qualité du groupe de sécurité (L = local, G = global, U = universel).

Ainsi, il est toujours possible de sélectionner rapidement les groupes de sécurité d'une unité:



ou dans un mode texte de connaître la qualité des groupes utilisée:

```
L:\sti-sg>cacls sti-it
L:\sti-sg\sti-it BUILTIN\
Administrators:(OI)(CI)F
         STI\stiit-AdminG:(OI)(CI)F
         STI\stiit-AllU:(OI)(CI)R
         STI\stiit-StaffG:(OI)(CI)R
```

Pour respecter les limites techniques du NAS, un espace disque est créé par institut, un partage de fichier est créé sous

son nom. Les laboratoires partagent cet espace commun.

Un lien DFS permet d'accéder directement au laboratoire, pour les usagers dont les ordinateurs Windows sont dans le domaine.

Une configuration personnalisée

Pour permettre un réglage conforme au désir des utilisateurs, deux éléments sont introduits:

- Une allocation extraordinaire d'équivalent usager. Ce mécanisme permet de répondre aux demandes quotidiennes où les exceptions sont plus fréquentes que la règle. Si une allocation uniforme de 3 Go par usager répond à une configuration standard, la possibilité d'augmenter l'allocation pour tenir compte d'historique ou de circonstances inhabituelles permet de satisfaire la quasi-totalité des usagers.
- Une répartition entre espaces privés et publics.

Ces deux paramètres sont introduits à la création d'une unité, ils peuvent être modifiés dès que le besoin se fait sentir.

Des quotas modulables

L'espace alloué à un laboratoire est défini. Seule, l'arrivée de nouveaux collaborateurs augmente sa taille.

Un principe de vase communicant est mis en place. Tout ajout ou suppression de quota chez un usager entraîne l'effet inverse pour le compte de l'administrateur. Pour éviter que des fichiers mis en commun par un usager soient décomptés, l'administrateur peut utiliser son quota. Il reste à traiter trois exceptions:

1. **Un usager qui collabore à un laboratoire extérieur**. Si un usager collabore avec un laboratoire qui ne fait pas partie de son institut, il obtient automatiquement un quota de 300 Mo dans celui-ci.
2. **Un compte de service *parasite***. Pour éviter la prolifération de comptes parasites bénéficiaires de l'espace équivalent à un passager, soit 3 Go, un usager non référencé obtient un espace sans conséquence sur l'espace global, soit 20 Mo.
3. **Un espace d'échange**. Il est parfois nécessaire de disposer d'un lieu pour déposer des fichiers à traiter, par exemple pour la commande de service. L'espace alloué par usager est convenable, par exemple 300 Mo, mais la taille de l'espace de stockage est limitée, par exemple 3 Go avec uniquement la possibilité de déposer un fichier.

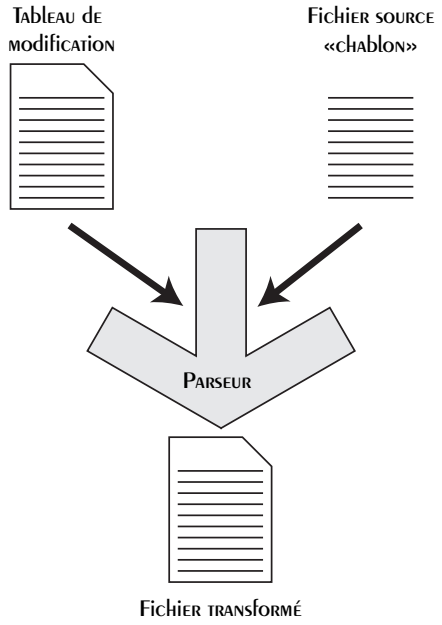
Adldap2 dans le détail

La capacité de gérer le NAS depuis un poste Windows en utilisant un API standard représente un vrai tour de force technique de la part d'EMC[2]. L'adaptation du programme existant consiste à:

- générer la modification de la structure du NAS comme suite à l'ajout d'un laboratoire, et conserver sa configuration spécifique;
- adapter les flux de données de création, déplacement et suppression d'usager avec des unités présentes dans le NAS;
- créer et gérer les quotas des usagers;
- gérer un fichier de configuration propre au domaine;
- être capable de s'intégrer dans un flux de traitement extérieur, en particulier la capacité de travailler récursivement pour réaliser l'ensemble des changements;
- accepter un mode manuel, ou automatique avec une limite aux nombres de modifications apportées.

L'ensemble des modifications est réalisé et **Adldap**[2] se trouve en production dans 3 domaines depuis janvier 2005 sans incident, représentant 1880 usagers et 247 unités. Les seuls arrêts observés ont été occasionnés par la présence de garde-fous internes évitant de traiter des données dont l'organisation a été modifiée ou dont la structure est incohérente, comme la présence de doublon.

La génération des modifications de la configuration du NAS utilise un mécanisme identique de transformation utilisée pour les autres fonctions:



Cette démarche est particulièrement efficace, remplaçant la réécriture de programme réalisant une tache définie par une simple modification d'un fichier texte.

Ainsi, le passage du format utilisé par l'importation d'usagers dans le NAS avec SecurityMapper par l'utilisation du UserMapper nécessite uniquement la modification d'une

Chablon UserMapper	`'Template to generate NAS-EMC group`
	`'`
	`' Version 1.0 , 4 janvier 2005`
	`' Version 1.1 , 12 janvier 2005, version UserMapper`
	`'`
	`' format to be`
	`'SID:*:GID:domain sti.intranet.epfl.ch`
	`##SecGroupSID##:*:##SecGroupGID##:domain sti.intranet.epfl.ch`

Chablon SecurityMapper	`'Template to generate NAS-EMC group`
	`'`
	`' Version 1.0 , 4 janvier 2005`
	`' Version 1.1 , 12 janvier 2005, version SecurityMapper`
	`'`
	`' format to be`
	`' <Nom groupe.domaine>:*:GID:`
	`##sAMaccount##.STI:*:##SecGroupGID##:`

ligne dans un fichier texte.

Un autre exemple est la création d'un laboratoire dans le NAS qui réalise également l'inscription du chemin réseau au niveau du DFS:

Chablon: NewNaslabo.tpl

```
'Template to create an unit in NAS system
'
' Version 1.0 , 30 janvier 2004
' Based of the same structure of current hierarchie generated by ADSI VBS Script
'
@rem cree dossier de l'unite ##Name##
MKDIR ##NASpath##\##Name##
@rem cree dossier commun
MKDIR ##NASpath##\##Name##\##NameStriped##-commun
...
@rem donne les droits
cacls ##NASpath##\##Name## /t /g Administrators:F ##NameStriped##-AdminG:F ##NameStriped##-StaffG:
R ##NameStriped##-AllU:R
@rem
@rem vbs work on NAS ("Microsoft.DiskQuota.1")
quotaadd ##NASdrive## ##NASadminQuota## admin##NameStriped##@sti.intranet.epfl.ch
@rem
....
@rem
@rem don't make new share (share are in institut level)
@rem
@rem make new DFS entry (in sub level of share for institut)
dfscmd /map \\sti\net\##Name## \\stisrv\##Institut##\##Name## "##Name## home in ##Institut##"
```

Résultat a exécuter: NewNasLabo_sti-sg_sti-it.cmd

```
@rem cree dossier de l'unite sti-it
MKDIR L:sti-sg\sti-it
@rem cree dossier commun
MKDIR L:sti-sg\sti-it\stiit-commun
...
@rem donne les droits
cacls L:sti-sg\sti-it /t /g Administrators:F stiit-AdminG:F stiit-StaffG:R stiit-AllU:R
@rem
@rem vbs work on NAS ("Microsoft.DiskQuota.1")
quotaadd L 32212254720 adminstiit@sti.intranet.epfl.ch
@rem
...
@rem
@rem don't make new share (share are in institut level)
@rem
@rem make new DFS entry (in sub level of share for institut)
dfscmd /map \\sti\net\sti-it \\stisrv\sti-sg\sti-it "sti-it home in sti-sg"
```

Une gestion indépendante d'une plate-forme

Dans un environnement hétérogène, il est indispensable d'offrir des interfaces sans relation avec la plate-forme technique. L'utilisation d'Inform décrit par Pierre Crevoisier dans le numéro précédent, ***http://dit.epfl.ch/publications-spip/article.php3?id_article=864,*** offre une interface simple pour créer des formulaires électroniques.

La présence sous-jacente de base de données permet une utilisation optimale.

Trois formulaires sont maintenant disponibles à l'adresse ***http://sti.epfl.ch/intranet/informatique/nas/***:

- la création d'une unité dans le NAS;
- la création d'utilisateur de service;
- la création et l'appartenance ont un groupe de sécurité.

Ces formulaires évitent à un administrateur dont l'environnement se compose exclusivement de Macintosh ou d'ordinateur sous Unix-Linux de devoir installer une machine virtuelle Windows connectée au domaine réel pour installer une console d'administration. Pour les unités où les ordinateurs se trouvent dans le domaine, une délégation de gestion s'impose généralement comme la solution la plus simple pour gérer son parc de PC Windows.

Un réglage fin des quotas réalisé par l'administrateur local

Dès la mise en place d'une gestion par quota, il apparaît vite la nécessité de disposer d'un rapport sur l'utilisation du disque et de pouvoir modifier le quota alloué à un usager. Chaque administrateur d'une unité possède l'accès d'un dossier contenant un rapport quotidien sur ses usagers. Le format utilisé peut être facilement importé depuis un tableur comme le montre la figure en page suivante.

La modification des quotas contenus dans ce fichier entraîne la modification des quotas alloués dans les limites établies. Une modification de quota réussie provoque l'archivage automatique de la consigne. Ainsi, un historique des modifications est conservé. Un développement en cours par deux apprentis dans notre Faculté, Claudia Mermoud STI-IT et Marc Longchamp I2S-LA permettra une visualisation et manipulation des quotas avec un programme réalisé en Java.

Un accès simple depuis les postes clients

Le premier objectif d'un système unifié est d'offrir une méthode simple de connexion d'un ordinateur depuis un poste client, quatre exemples:

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Faculté STI Sciences et Techniques de l'ingénieur

adldap2: Tableau de bord – Quotas de l'unités LSA1

Rapport du 06.05.2005 14.02

Unité lsa1

	Nb	NAS Attribué	Utilisé	
Unité	1	312.00 Go	58.98 Go	18.9%
Usager	36		58.98 Go	

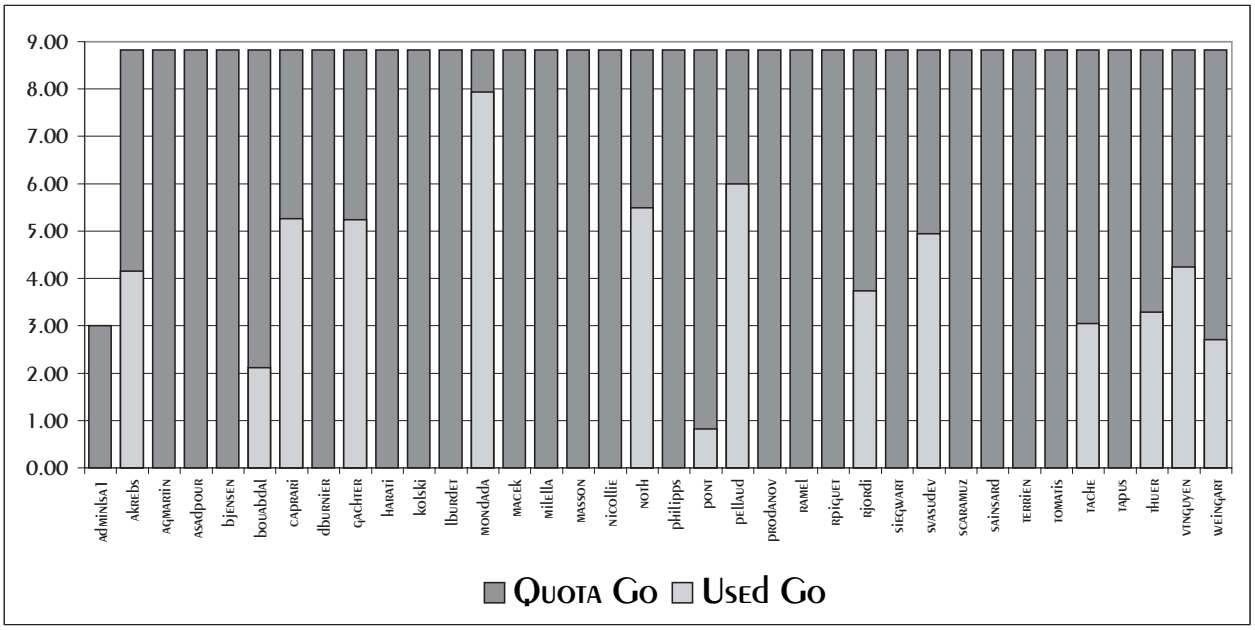


Tableau de bord d'Adldap^2: rapport sur l'utilisation du NAS pour une unité

Pour un Macintosh

- vérifier que l'accès SMB est activé et que le groupe de travail correspond au domaine (STI) avec le bon serveur Wins (`128.178.15.44`);
- vérification que les DNS sont ceux d'Active Directory (`128.178.15.227` & `128.178.15.228`);
- vérifier que le compte Active Directory est actif et que la synchronisation des mots de passe est correcte;
- réaliser la connexion `smb://stisrv/institut/laboratoire/nomimap`, institut et laboratoire sont l'acronyme de l'unité;
- pour le nom d'utilisateur, introduire votre nom `imap` et pour le mot de passe, votre mot de passe `gaspar`.

Pour un PC Linux avec l'aide de Rolf Jordi, STI-I2S-LSA1

- vérification que les DNS sont ceux d'Active Directory (`128.178.15.227` & `128.178.15.228`);
- vérifier que le compte Active Directory est actif et que la synchronisation des mots de passe est correcte;
- créer un point de montage `/mnt/sti`;
- monter le volume de l'institut `mount -t cifs -o user=nomimap,domain=sti //stisrv/institut /mnt/sti/`, institut et laboratoire sont l'acronyme de l'unité;
- un mot de passe sera demandé, introduire votre mot de passe Gaspar;
- aller à votre dossier personnel `/mnt/sti/laboratoire/nomimap`.

Pour un PC Windows hors domaine

- vérifier le bon serveur Wins (`128.178.15.44`);
- vérification que les DNS sont ceux d'Active Directory (`128.178.15.227` & `128.178.15.228`);
- vérifier que le compte Active Directory est actif et que la synchronisation des mots de passe est correcte
- réaliser la connexion avec l'outil **monter un disque**; choisir la lettre ou monter le volume, pour le nom de serveur `\\stisrv\institut`, institut et laboratoire sont l'acronyme de l'unité;
- pour le nom d'utilisateur, introduire votre nom `sti\nomimap`; pour le mot de passe, introduire votre mot de passe `Gaspar`;
- aller au bon dossier, `Lettre:\laboratoire\nomimap`.

Pour un PC Windows dans le domaine

- le serveur WINS doit déjà être correct (`128.178.15.44`);
- idem pour le DNS (`128.178.15.227` & `128.178.15.228`);
- idem pour le compte Active Directory;
- connectez-vous dans le domaine;
- Réaliser la connexion avec l'outil **monter un disque**; choisir la lettre ou monter le volume pour le nom de serveur `\\sti\net\laboratoire\nomimap`, laboratoire est l'acronyme de l'unité;
- le mot de passe a déjà été introduit lors de l'ouverture de la session et le laboratoire possède un lien DFS.

En conclusion, une nouvelle méthode de travail

L'arrivée du système NAS offre la chance de repenser la place du stockage dans notre travail quotidien. La diversité des moyens offerts à un utilisateur pour stocker des informations est très élevée. Il peut paraître curieux d'offrir un espace centralisé alors qu'une clé USB de 0.5 ou 1 Go offre un bon rapport capacité – coût – autonomie.

Au contraire, la persistance des données du NAS doit être utilisée pour conserver les informations les plus pertinentes. Sa capacité d'augmenter la taille des données sans incidence sur sa gestion et la conservation à court terme des derniers états d'un fichier offre une flexibilité élevée. Une utilisation naïve serait de l'utiliser comme un espace de sauvegarde, un super coffre-fort conservant dans le cas le plus absurde un gros fichier archivé chaque jour représentant la sauvegarde de plusieurs milliers de documents. Il doit être utilisé comme l'espace central de conservation avec un accès par le réseau qui offre le maximum de flexibilité.

La migration de services, une démarche darwinienne ?

À la transmission des caractéristiques biologiques par la sélection génétique des espèces, les humains ont substitué une sélection des idées permettant à l'humanité des progrès importants.

Cette transmission du savoir peut prendre une forme spectaculaire. Le peuple ouïgour aux confins de l'Asie centrale n'a pas hésité d'utiliser l'écriture du peuple dominant, passant successivement par l'écriture runique, sogdiane, arabe, chinoise et finalement cyrillique sur plus de 2000 ans. Le paradoxe tient au fait que les trois dernières écritures soient restées vivaces, offrant la vision surréaliste d'un hebdomadaire véhiculant la même langue écrite dans trois écritures totalement différentes dans le sens de la lecture et de l'alphabet. Un parallèle peut être établi avec le NAS qui à partir d'une base Unix offre des services intégrés dans le monde Windows.

La capacité de l'entreprise EMC^2 de gérer l'ensemble de ce processus est particulièrement intéressante. Par exemple l'outil de migration (CDMS) permet la migration *en ligne* de données, en clonant les attributs du serveur source comme le nom ou l'adresse IP voir l'adresse MAC sans arrêter l'accès des données par les usagers. Un équipement équivalant au nôtre a permis aux Hôpitaux de Paris de migrer ainsi plus de 450 serveurs répartis sur 18 sites vers 2 sites.

Une extension possible du NAS serait son couplage avec un système de gestion hiérarchique des données sur la durée (HSM). Ce système offrira la possibilité de créer un vrai espace de stockage illimité où les données sont conservées sur le support le plus adapté en fonction de leur fréquence de modification: du disque dur rapide à une technologie plus lente pour être finalement archivés sur des disques optiques.

Face aux changements apportés par ce genre de techniques, certains administrateurs système pourraient avoir une réaction comme le Luddisme (révolte ouvrière devant la montée en puissance de l'automatisation du travail au début de la révolution industrielle au 18e) .

Mais si un serveur de fichier local perd son rôle central, il retrouve une seconde vie en tant qu'espace intermédiaire de stockage de données.

De manière similaire, la consolidation des serveurs d'applications est offerte par le serveur VMWare ESX. Cet outil permet de faire cohabiter des systèmes d'exploitation différents sur la même machine physique en utilisant des machines virtuelles.

Pour faciliter le transfert, un outil (P2V de VMWare) permet de cloner un serveur physique pour le transformer en machine virtuelle.

Enfin, la notion d'une machine virtuelle dans un ordinateur puissant nous ramène au système d'exploitation MVS-VM d'IBM. L'informatique étant, comme on sait, un éternel recommencement ! ■